

(51)Int.Cl.⁵

G 10 L 7/08

識別記号 厅内整理番号

A 8842-5H

F I

技術表示箇所

審査請求 未請求 請求項の数2(全11頁)

(21)出願番号 特願平4-42394

(22)出願日 平成4年(1992)1月31日

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 赤羽 誠

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(72)発明者 西口 正之

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

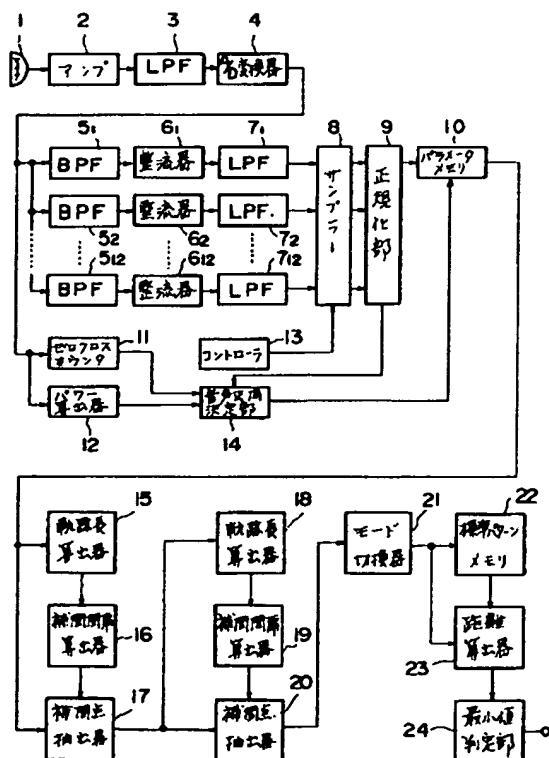
(74)代理人 弁理士 稲本 義雄 (外1名)

(54)【発明の名称】 音声認識装置

(57)【要約】

【目的】 音声の認識率を向上させる。

【構成】 サンプラー8において、各チャンネル(BPF(バンドパスフィルタ)5_i、整流回路6_iおよびLPF(ローパスフィルタ)7_iから成る回路)より出力されるパワーが、例えばサンプリング周波数8kHz(周期0.125ms)でサンプリングされる。そして、そのサンプリングされた各パワーに、各チャンネルの帯域幅(BPF5₁乃至5₁₂の帯域幅)の逆数に比例する値に最も近い例えば64の倍数の時間窓長を有する時間窓が掛られ、時間窓がその時間窓長の1/2ずつシフトされながら、その時間窓区間のパワーの総和が、特徴パラメータとして算出される。



【特許請求の範囲】

【請求項1】 音声信号を所定の周波数帯域に分割し、各帯域ごとに特徴パラメータを抽出する特徴パラメータ抽出手段と、

前記帯域の周波数に対応して、前記特徴パラメータ抽出手段により特徴パラメータが抽出される時間窓長を制御する窓長制御手段と、

前記音声信号の標準パターンを記憶する標準パターン記憶手段と、

前記窓長制御手段により制御された時間窓長ごとに出力される特徴パラメータと、前記標準パターン記憶手段に記憶された標準パターンとの距離を算出する距離算出手段と、

前記距離算出手段により算出された距離により前記音声信号を認識する認識手段とを備えることを特徴とする音声認識装置。

【請求項2】 前記窓長制御手段は前記音声信号を帯域分割する帯域幅に逆比例するように前記時間窓長を制御することを特徴とする請求項1に記載の音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声を認識する場合に用いて好適な音声認識装置に関する。

【0002】

【従来の技術】 図5は、従来の音声認識装置の一例の構成を示すブロック図である。マイク1は、入力された音声（例えば単語）を電気信号である音声信号に変換する。音響分析部31は、マイク1より出力された音声信号を音響分析し、音声の特徴パラメータを抽出する。モード切換器21は、装置のモードが登録モードである場合、音響分析部31より出力される特徴パラメータを標準パターンメモリ22に供給し、装置のモードが認識モードである場合、音響分析部31より出力される特徴パラメータを入力パターンメモリ32に供給する。

【0003】 標準パターンメモリ22は、登録モード時にモード切換器21より供給される特徴パラメータを標準パターンとして記憶する。入力パターンメモリ32は、認識モード時にモード切換器21より供給される特徴パラメータを入力パターンとして一時記憶する。距離算出器23は、標準パターンと入力パターンとの距離を計算する。即ち、距離算出器23は、標準パターンメモリ22に記憶された標準パターンとしての特徴パラメータと、入力パターンメモリ32に記憶された入力パターンとしての特徴パラメータとの距離を計算する。最小値判定部24は、距離算出器23により計算された標準パターンと入力パターンとの距離の中から、入力パターンとの距離を最も短くする標準パターンに対応する単語を認識結果として出力する。

【0004】 このように構成される音声認識装置では、モードが登録モードである場合、マイク1に入力された

音声が、音響分析部31において音響分析され、特徴パラメータが抽出される。特徴パラメータは、モード切換器21を介して標準パターンメモリ22に標準パターンとして記憶される。

【0005】 モードが認識モードである場合、マイク1に入力された音声が、音響分析部31において音響分析され、特徴パラメータが抽出される。特徴パラメータは、モード切換器21を介して入力パターンメモリ32に入力パターンとして一時記憶される。距離算出器23において、あらかじめ標準パターンメモリ22に記憶されたすべての標準パターンと、入力パターンメモリ32に記憶された入力パターンとの距離がそれぞれ計算される。最小値判定部24において、距離算出器23により計算された標準パターンと入力パターンとの距離の中から、入力パターンとの距離を最も短くする標準パターンが選択され、そのパターンに対応する単語が認識結果として出力される。

【0006】

【発明が解決しようとする課題】 ところで、音響分析部31で特徴パラメータを抽出するための音響分析方法は種々考えられるが、例えばフィルタバンクによる手法（フィルタバンク法）がある。

【0007】 即ち、BPF（バンドパスフィルタ）と整流回路とから成る回路を1チャンネルとし、そのBPFの通過帯域を、例えば周波数軸上（対数軸上）で等間隔に分割して複数個並べ、各チャンネルごとの（帯域ごとの）、例えばパワーなどのスペクトルパターンの時間変化を表す特徴パラメータを抽出する方法である。

【0008】 このようなフィルタバンクの設計例を図6に示す。図6のフィルタバンクは、340Hz乃至3400Hzの帯域を、BPFの通過帯域を周波数軸上（対数軸上）で等間隔に12CH（12チャンネル）に分割するフィルタバンクである。このようなフィルタバンクにおいて、音声信号が例えば8kHzのサンプリング周波数でサンプリングされ、図7に示すように128サンプル（16ms）の時間窓長で、且つ64サンプル（8ms）ごとに特徴パラメータとしてのパワーが算出される場合、各チャンネルより出力される1秒あたりのデータ量（数）は、

$$40 \quad 1(s) / 8(ms) = 125(\text{個}/s)$$

となる。よって、12チャンネルのフィルタバンクにおいては、上記データ量を12倍して、

$$125(\text{個}/s) \times 12(\text{チャンネル}) = 1500(\text{個}/s)$$

となり、フィルタバンクのチャンネル数に比例してデータ量が増え、計算量が増加する課題があった。

【0009】 さらに、このような方法においては、各帯域ごとの特徴パラメータとしてのパワーが、8ms（64サンプル）などの所定の一定時間ごとに抽出され（図8）、以後認識処理が行われる。しかしながら、人間の

聴覚においては、低域ほど音響分析に時間をかける、即ち高域ほど短い時間で音響分析を行っていることが知られているが、従来の装置では、低域も高域も同じ時間だけ分析するようにしているため、実際の人間の聴覚に近い分析を行うことができず、正確に音声認識をすることが困難である課題があった。

【0010】本発明は、このような状況に鑑みてなされたものであり、人の聴覚の時間分解能を考慮し、音声の認識率を向上させるものである。

【0011】

【課題を解決するための手段】請求項1に記載の音声認識装置は、音声信号を所定の周波数帯域に分割し、各帯域ごとに特徴パラメータを抽出する特徴パラメータ抽出手段としてのBPF（バンドパスフィルタ）5₁乃至5₁₂、整流器6₁乃至6₁₂、LPF（ローパスフィルタ）7₁乃至7₁₂、およびサンプラー8と、帯域の周波数に対応して、BPF5₁乃至5₁₂、整流器6₁乃至6₁₂、LPF7₁乃至7₁₂、およびサンプラー8により特徴パラメータが抽出される時間窓長を制御する窓長制御手段としてのコントローラ13と、音声信号の標準パターンを記憶する標準パターン記憶手段としての標準パターンメモリ22と、コントローラ13により制御された時間窓長ごとに出力される特徴パラメータと、標準パターンメモリ22に記憶された標準パターンとの距離を算出する距離算出手段としての距離算出器23と、距離算出器23により算出された距離により音声信号を認識する認識手段としての最小値判定部24とを備えることを特徴とする。

【0012】この音声認識装置は、コントローラ13に音声信号を帯域分割する帯域幅に逆比例するように時間窓長を制御させることができる。

【0013】

【作用】請求項1に記載の音声認識装置においては、音声信号を所定の周波数帯域に分割し、各帯域の周波数に対応して特徴パラメータを抽出する時間窓長を制御する。そして、標準パターンメモリ22に記憶されている標準パターンと、制御された時間窓長ごとに出力される特徴パラメータとの距離を算出し、算出された距離により音声信号を認識する。従って、人の聴覚の時間分解能に合った時間間隔で特徴パラメータを得るようにすることができるので、音声の認識率を向上させることができ。また、距離を算出するときの計算量を減少させることができる。

【0014】コントローラ13に音声信号を帯域分割する帯域幅に逆比例するように時間窓長を制御させることができの場合においては、一般的に広帯域に帯域分割される音声信号の高域の変化に対応することができるのと、さらに音声の認識率を向上させることができる。

【0015】

【実施例】図1は、本発明の音声認識装置の一実施例の

10

20

30

40

構成を示すブロック図である。図5における場合と対応する部分には、同一の符号を付してある。アンプ2は、マイク1より出力される音声信号を増幅する。A/D変換器4は、LPF3を介してアンプ2より供給される音声信号を例えば8kHzのサンプリング周波数でサンプリングする。LPF（ローパスフィルタ）3は、アンプ2より出力される音声信号の周波数帯域を、次段のA/D変換器4におけるサンプリング周波数の1/2の周波数（ナイキストレート）、即ち4kHzに帯域制限する。

【0016】BPF（バンドパスフィルタ）5₁乃至5₁₂は、図2に示すように、340Hz乃至3400Hzの周波数帯域を対数軸上で等間隔に分割した例えば12の帯域を、それぞれの通過域として有するディジタルフィルタであり、A/D変換器4より供給されるディジタル音声信号を、各フィルタ（BPF5₁乃至5₁₂）の通過域にしたがって通過させる。なお、図2のCH（チャンネル）番号は、図1のBPF5₁乃至5₁₂のサフィックスに対応する。

【0017】整流回路6₁乃至6₁₂は、BPF5₁乃至5₁₂より出力される各帯域ごとの音声信号をそれぞれ整流し、LPF（ローパスフィルタ）7₁乃至7₁₂に供給する。LPF7₁乃至7₁₂は、整流器6₁乃至6₁₂より供給される、整流された音声信号を平滑化し、BPF5₁乃至5₁₂の各通過域における音声信号のパワーとしてサンプラー8に出力する。

【0018】ここで、BPF5_i、整流回路6_iおよびLPF7_iとから成る回路を1チャンネルとし、以下第iチャンネルと呼ぶ（1≤i≤12）。

【0019】サンプラー8は、各チャンネルより出力されるパワーを、例えばサンプリング周波数8kHz（周期0.125ms）でサンプリングする。さらに、サンプラー8は、コントローラ13に制御され、サンプリングした各パワーに、各チャンネルの帯域幅（BPF5₁乃至5₁₂の帯域幅）の逆数に比例する時間窓長を有する時間窓を掛け、時間窓をその時間窓長の1/2ずつシフトしながら、その時間窓区間のパワーの総和を算出する。即ち、サンプラー8は、例えば図2に示すように、第12チャンネルの帯域幅（通過域が2806Hz乃至3400Hz）594Hzの逆数より、BPF5₁乃至5₁₂の帯域幅の逆数を正規化した帯域幅の逆数比に比例する時間窓長を有する時間窓を各チャンネルより出力されるパワーに掛け、時間窓をその時間窓長の1/2ずつシフトしながら（例えば第12チャンネルにおいては、32サンプル（4ms）ずつシフトしながら）、その時間窓区間のパワーの総和を算出する。

【0020】但し、本実施例においては、装置を簡単に構成するため、時間窓長は各チャンネルの帯域幅の逆数比に最も近い64の倍数になるように設定してある（図2）。従って、上述したように、第12チャンネルの時

間窓長を64サンプル(8ms)とした場合、第11チャンネルにおいては、その帯域幅の逆数比が、1.2122であるから、時間窓長は、本来この逆数比に比例させて77(=64×1.2122)サンプルとなるが、本実施例では、この値に最も近い64の倍数である64サンプルに設定してある。第10チャンネルにおいては、その帯域幅の逆数比が、1.4703であるから、時間窓長は、本来この逆数比に比例させて96(=64×1.4703)サンプルとなるが、本実施例では、この値に最も近い64の倍数である64サンプルに設定してある。以下同様にして、図2のように各チャンネルの時間窓長を設定してある。

【0021】正規化部9は、サンプラー8より出力される各チャンネルのパワーを正規化し、パラメータメモリ10および音声区間決定部14に供給する。

【0022】ゼロクロスカウンタ11は、A/D変換器4より出力された音声信号の所定のフレーム周期T内のゼロクロス数をカウントし、音声区間決定部14に供給する。パワー算出器12は、A/D変換器4より出力された音声信号のフレーム周期T内のパワーを算出し、音声区間決定部14に供給する。音声区間決定部14は、正規化部9より供給される各チャンネルのパワー(正規化されたパワー)、ゼロクロスカウンタ11より供給されるゼロクロス数、およびパワー算出器12より供給される音声信号のパワーから、現在のフレームn(分析区間)が、無音、無声音、または有声音のうちのどの状態であるかを判定する。

【0023】パラメータメモリ10は、音声区間決定部14で現在のフレームnが、無音区間でない音声(無聲音または有声音)区間であると判定された場合、正規化部9より出力される現在のフレームnにおける、各チャンネル(第1乃至第12チャンネル)のパワーP

$$p = SL / T_1 + 1$$

である。

【0026】軌跡長算出器18は、軌跡長算出器15と同様に、補間点抽出器17より出力される新たな特徴パラメータ系列R_i(p)が、そのパラメータ空間内に描く直線近似による軌跡長SL'を算出する。補間間隔算出器19は、パラメータ空間内に描かれた特徴パラメータ

$$T_1' = SL' / (M - 1)$$

【0027】補間点抽出器20は、補間点抽出器17と同様に、特徴パラメータR_i(p)がパラメータ空間内に描く軌跡を直線近似した軌跡に沿って、補間間隔算出器19により算出された再サンプリング間隔T₁'で再サンプリングし、さらに新たな特徴パラメータ系列Q_i(m)(1≤m≤M)を出力する。

【0028】コントローラ13は、サンプラー8が、各チャンネルより出力されるパワーをサンプリングし、その各パワーに、各チャンネルの帯域幅(BPF5₁乃至5₁₂の帯域幅)の逆数に比例する時間窓長を有する時間

_i(n)(1≤i≤12)の記憶を開始する。さらに、パラメータメモリ10は、音声区間決定部14で現在のフレームnが、音声区間から無音区間になったと判定された場合、正規化部9より出力される現在のフレームnにおける、各チャンネル(第1乃至第12チャンネル)のパワーP_i(n)(1≤i≤12)の記憶を終了する。従って、パラメータメモリ10は、音声区間Nにおける各チャンネルのパワーP_i(n)(1≤i≤12, n=1, 2, ..., N)を記憶する(Nは、音声区間の長さに対応する)。

【0024】軌跡長算出器15は、パラメータメモリ10に記憶された音声の特徴パラメータとしてのパワーP_i(n)が、そのパラメータ空間内に描く直線近似による軌跡長SLを算出する。即ち、軌跡長算出器15は、例えば2チャンネルのパワーP₁(n)およびP₂(n)のみを考えた場合、この2つのパラメータ(2チャンネルのパワーP₁(n)およびP₂(n))が、図3に示すような2次元のパラメータ空間(2次元平面)に描く軌跡長を算出する(パラメータの軌跡を、図中×印および●印で示す(●印は、軌跡の始点および終点を示す))。

【0025】補間間隔算出器16は、パラメータ空間内に描かれたパワーP_i(n)の軌跡を再サンプリングするための間隔(再サンプリング間隔)T₁を、軌跡長算出器15により算出された軌跡長SLをもとに算出する。補間点抽出器17は、パラメータメモリ10に記憶された音声の特徴パラメータとしてのパワーP_i(n)がパラメータ空間内に描く軌跡を直線近似した軌跡に沿って、図3に○印で示すように、補間間隔算出器16により算出された再サンプリング間隔T₁で、パワーP_i(n)の軌跡を再サンプリングし、新たな特徴パラメータ系列R_i(p)を出力する。但し、

(1)

タR_i(p)の軌跡を、さらに再サンプリングするための間隔(再サンプリング間隔)T₁を、あらかじめ設定された再サンプリング点数M、および軌跡長算出器18により算出された軌跡長SL'から、次式にしたがって算出する。

(2)

窓を掛け、時間窓をその時間窓長の1/2ずつシフトしながら、その時間窓区間のパワーの総和を算出するように制御する。

【0029】次に、その動作について説明する。マイク1より入力された音声は、アンプ2およびLPF3を介してA/D変換器4に供給され、A/D変換器4において、8kHzのサンプリング周波数でサンプリングされ、BPF5₁乃至5₁₂に出力される。

【0030】BPF(バンドパスフィルタ)5₁乃至5₁₂において、図2に示すように、340Hz乃至340

0 Hz の帯域を対数軸上で等間隔に分割した 12 の帯域（通過域）にしたがって、A/D 変換器 4 より出力された音声信号がそれぞれ通過し、整流器 6₁ 乃至 6₁₂ に出力される。

【0031】整流回路 6₁ 乃至 6₁₂ において、BPF 5₁ 乃至 5₁₂ より出力される各帯域ごとの音声信号がそれぞれ整流され、LPF（ローパスフィルタ）7₁ 乃至 7₁₂ において、平滑化され、BPF 5₁ 乃至 5₁₂ の各通過域における音声信号のパワーとしてサンプラー 8 に出力される。

【0032】サンプラー 8 において、各チャンネル（BPF 5_i、整流回路 6_i および LPF 7_i から成る回路）より出力されるパワーが、サンプリング周波数 8 kHz（周期 0.125 ms）でサンプリングされる。そして、そのサンプリングされた各パワーに、各チャンネルの帯域幅（BPF 5₁ 乃至 5₁₂ の帯域幅）の逆数に比例する値に最も近い 64 の倍数の時間窓長を有する時間窓が掛けられ、時間窓がその時間窓長の 1/2 ずつシフトされながら、その時間窓区間のパワーの総和が算出される。

【0033】即ち、サンプラー 8 において、図 2 に示すように、第 1 チャンネルの帯域幅の逆数比 8.2500 に比例する値で、64 の倍数に最も近い時間窓長 512 サンプル（64 ms）の時間窓が、第 1 チャンネルより出力されるパワーに掛けられ、その時間窓が、時間窓長の 1/2 周期（パワー出力周期）、即ち 256 サンプル（32 ms）だけシフトされながら（図 4）、その時間窓区間のパワーの総和が出力される。第 2 チャンネルにおいては、その帯域幅の逆数比 6.8276 に比例する値で、64 の倍数に最も近い時間窓長 448 サンプル（56 ms）の時間窓が、第 2 チャンネルより出力されるパワーに掛けられ、その時間窓が、時間窓長の 1/2 周期（パワー出力周期）、即ち 224 サンプル（28 ms）だけシフトされながら（図 4）、その時間窓区間のパワーの総和がサンプラー 8 より出力される。

【0034】以下同様にして、サンプラー 8 において、図 2 に示す時間窓区間内の各チャンネルのパワーの総和が、パワー出力周期だけシフトされながら出力される。

【0035】ここで、図 2 の最右列のデータレートは、各チャンネルより出力される 1 秒あたりのデータ量（数）を示し、さらにその最下段は、全チャンネルより出力されるデータ量（数）を示している。前述した、図 7 の場合（128 サンプル（16 ms）の時間窓長で、且つ 64 サンプル（8 ms）ごとに特徴パラメータとしてのパワーが算出される場合）のデータ量（1 秒あたり 1500 データ）と比較して、図 2 の場合の方がデータ量が少なく（1 秒あたり約 1388 データ）、従って、以下のブロックで行われる計算量も軽減することができる判ることが判る。

【0036】正規化部 9 において、サンプラー 8 より出力された各チャンネルのパワーが正規化され、パラメータメモリ 10 および音声区間決定部 14 に供給される。

【0037】一方、ゼロクロスカウンタ 11 において、A/D 変換器 4 より出力された音声信号の所定のフレーム周期 T 内のゼロクロス数がカウントされ、音声区間決定部 14 に供給される。パワー算出器 12 において、A/D 変換器 4 より出力された音声信号のフレーム周期 T 内のパワーが算出され、音声区間決定部 14 に供給される。音声区間決定部 14 において、正規化部 9 より供給される各チャンネルのパワー（正規化されたパワー）、ゼロクロスカウンタ 11 より供給されるゼロクロス数、およびパワー算出器 12 より供給される音声信号のパワーから、現在のフレーム n（分析区間）が、無音、無聲音、または有聲音のうちのどの状態であるかが判定される。

【0038】音声区間決定部 14 で現在のフレーム n が、無音区間でない音声（無聲音または有聲音）区間であると判定された場合、パラメータメモリ 10 において、正規化部 9 より出力される現在のフレーム n における、各チャンネル（第 1 乃至第 12 チャンネル）のパワー $P_i(n)$ ($1 \leq i \leq 12$) の記憶が開始される。さらに、音声区間決定部 14 で現在のフレーム n が、音声区間から無音区間になったと判定された場合、パラメータメモリ 10 において、正規化部 9 より出力される現在のフレーム n における、各チャンネル（第 1 乃至第 12 チャンネル）のパワー $P_i(n)$ ($1 \leq i \leq 12$) の記憶が終了される。従って、パラメータメモリ 10 においては、音声区間 N における各チャンネルのパワー $P_i(n)$ ($1 \leq i \leq 12, n = 1, 2, \dots, N$) が記憶される（N は、音声区間の長さに対応する）。

【0039】軌跡長算出器 15 において、パラメータメモリ 10 に記憶された音声の特徴パラメータとしてのパワー $P_i(n)$ がパラメータ空間内に描く軌跡長（直線近似による軌跡長） S_L が算出される。

【0040】補間間隔算出器 16 において、パラメータ空間内に描かれたパワー $P_i(n)$ の軌跡を再サンプリングするための間隔（再サンプリング間隔） T₁ が、軌跡長算出器 15 により算出された軌跡長 S_L をもとに算出される。補間点抽出器 17 において、パワー $P_i(n)$ がパラメータ空間内に描く軌跡を直線近似した軌跡に沿って、補間間隔算出器 16 により算出された再サンプリング間隔 T₁ で再サンプリングされ、新たな特徴パラメータ系列 R_{i(p)} が、軌跡長算出器 18 および補間点抽出器 20 に供給される（p は、式（1）により決定される）。

【0041】軌跡長算出器 18 において、補間点抽出器 17 より出力される新たな特徴パラメータ系列 R_{i(p)} がパラメータ空間内に描く軌跡長（直線近似による軌跡長） S_L が算出される。補間間隔算出器 19

において、パラメータ空間内に描かれた特徴パラメータ $R_i(p)$ の軌跡を、さらに再サンプリングするための間隔（再サンプリング間隔） $T_{1'}$ が、あらかじめ設定された再サンプリング点数 M 、および軌跡長算出器 18 により算出された軌跡長 SL' から、式 (2) にしたがって算出される。

【0042】補間点抽出器 20 において、新たな特徴パラメータ $R_i(p)$ がパラメータ空間内に描く軌跡を直線近似した軌跡に沿って、補間間隔算出器 19 により算出された再サンプリング間隔 $T_{1'}$ で、新たな特徴パラメータ $R_i(p)$ が再サンプリングされ、さらに新たな特徴パラメータ系列 $Q_i(m)$ ($1 \leq m \leq M$) が出力される。

【0043】このように、軌跡長算出器 15、補間間隔算出器 16 および補間点抽出器 17 によりパラメータ空間内に描かれた特徴パラメータとしてのパワー $P_i(n)$ の軌跡が、その軌跡長に対応して再サンプリングされ、新たな特徴パラメータ $R_i(p)$ が算出された後、サンプリング点数を一定値 M にするために、さらに軌跡長算出器 15、補間間隔算出器 16 および補間点抽出器 17 により新たな特徴パラメータ $R_i(p)$ の軌跡が、その軌跡長に対応して再サンプリングされ、さらに新たな特徴パラメータ $Q_i(m)$ が算出され、モード切換器 21 に供給される。

【0044】装置のモードが登録モードである場合、モード切換器 21 において、補間点抽出器 20 より供給された特徴パラメータ $Q_i(m)$ が、標準パターンメモリ 22 に標準パターンとして記憶される。

【0045】装置のモードが認識モードである場合、モード切換器 21 において、補間点抽出器 20 より供給された特徴パラメータ $Q_i(m)$ が、入力パターンとして距離算出器 23 に出力される。距離算出器 23 において、あらかじめ標準パターンメモリ 22 に記憶されたすべての標準パターンと、モード切換器 21 より出力された入力パターンとしての特徴パラメータ $Q_i(m)$ との距離がそれぞれ計算され、最小値判定部 24 に供給される。最小値判定部 24 において、距離算出器 23 により計算された標準パターンと入力パターンとの距離の中から、入力パターンとの距離を最も短くする標準パターンが選択され、そのパターンに対応する単語が認識結果として出力される。

【0046】以上説明したように、人間の聴覚における時間分解能を考慮し、各チャンネルの周波数（帯域幅）に対応して、音声の特徴パラメータを抽出する時間窓長を設定するようにしたので、音声の認識率を向上させることができるだけでなく、装置の各ブロックにおける計算量を減少させることができる。

【0047】

【発明の効果】請求項 1 に記載の音声認識装置によれば、音声信号を所定の周波数帯域に分割し、各帯域の周

波数に対応して特徴パラメータを抽出する時間窓長を制御する。そして、標準パターン記憶手段に記憶されている標準パターンと、制御された時間窓長ごとに出力される特徴パラメータとの距離を算出し、算出された距離により音声信号を認識する。従って、人の聴覚の時間分解能に合った時間間隔で特徴パラメータを得るようにすることができるので、音声の認識率を向上させることができる。また、距離を算出するときの計算量を減少させることができる。

【0048】請求項 2 に記載の音声認識装置によれば、窓長制御手段に音声信号を帯域分割する帯域幅に逆比例するように時間窓長を制御させる。従って、一般的に広帯域に帯域分割される音声信号の高域の変化に対応することができるので、さらに音声の認識率を向上させることができる。

【図面の簡単な説明】

【図 1】本発明の音声認識装置の一実施例の構成を示すブロック図である。

【図 2】図 1 のバンドパスフィルタ 5₁ 乃至 5₁₂ の帯域幅と、サンプラー 8 より出力されるパワーに掛けられる時間窓長、並びにそのパワーが outputされる周期との関係を示す図である。

【図 3】サンプラー 8 より出力される特徴パラメータとしてのパワーの軌跡を示す平面図である。

【図 4】図 1 のバンドパスフィルタ 5₁ 乃至 5₁₂ よりサンプラー 8 を介して、特徴パラメータとしてのパワーが outputされる周期を示す図である。

【図 5】従来の音声認識装置の一例の構成を示すブロック図である。

【図 6】従来の音声認識装置におけるバンドパスフィルタの一設計例である。

【図 7】従来のフィルタバンクにおけるパワーを出力する周期およびパワーを算出する区間を説明する図である。

【図 8】従来のフィルタバンクにおけるパワーを出力する周期を説明する図である。

【符号の説明】

- 1 マイク
- 2 アンプ
- 3 L P F (ローパスフィルタ)
- 4 A/D 変換器
- 5₁ 乃至 5₁₂ B P F (バンドパスフィルタ)
- 6₁ 乃至 6₁₂ 整流器
- 7₁ 乃至 7₁₂ L P F (ローパスフィルタ)
- 8 サンプラー
- 9 正規化部
- 10 パラメータメモリ
- 11 ゼロクロスカウンタ
- 12 パワー算出器
- 13 コントローラ

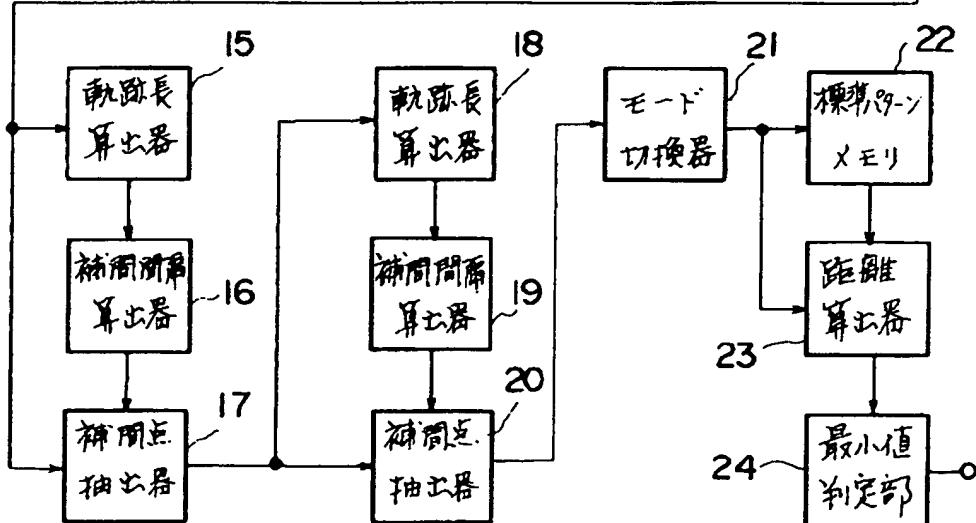
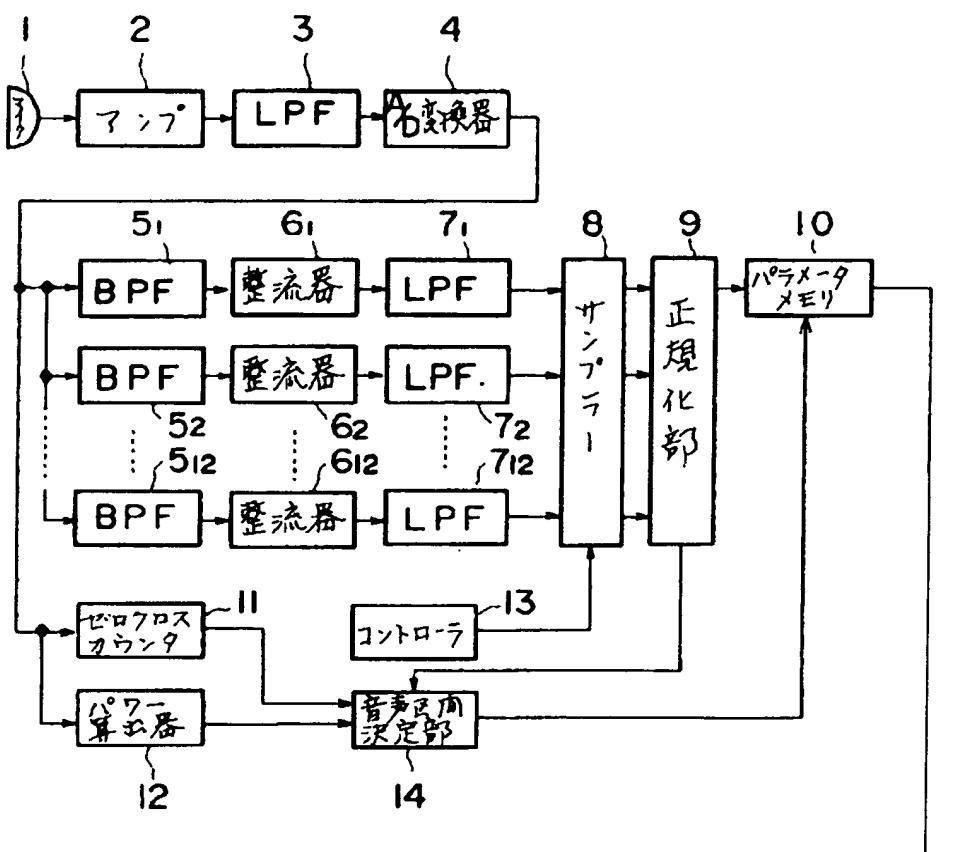
11

12

14 音声区間決定部
 15 軌跡長算出器
 16 補間間隔算出器
 17 補間点抽出器
 18 軌跡長算出器
 19 補間間隔算出器
 20 補間点抽出器

21 モード切換器
 22 標準パターンメモリ
 23 距離算出器
 24 最小値判定部
 31 音響分析部
 32 入力パターンメモリ

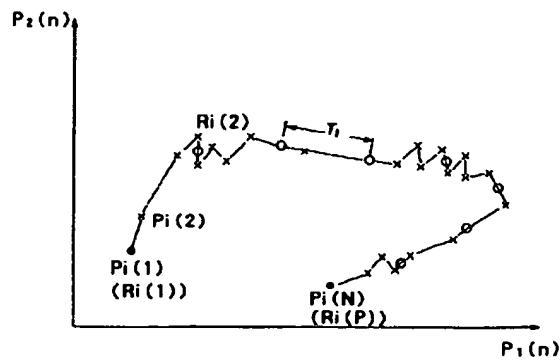
【図1】



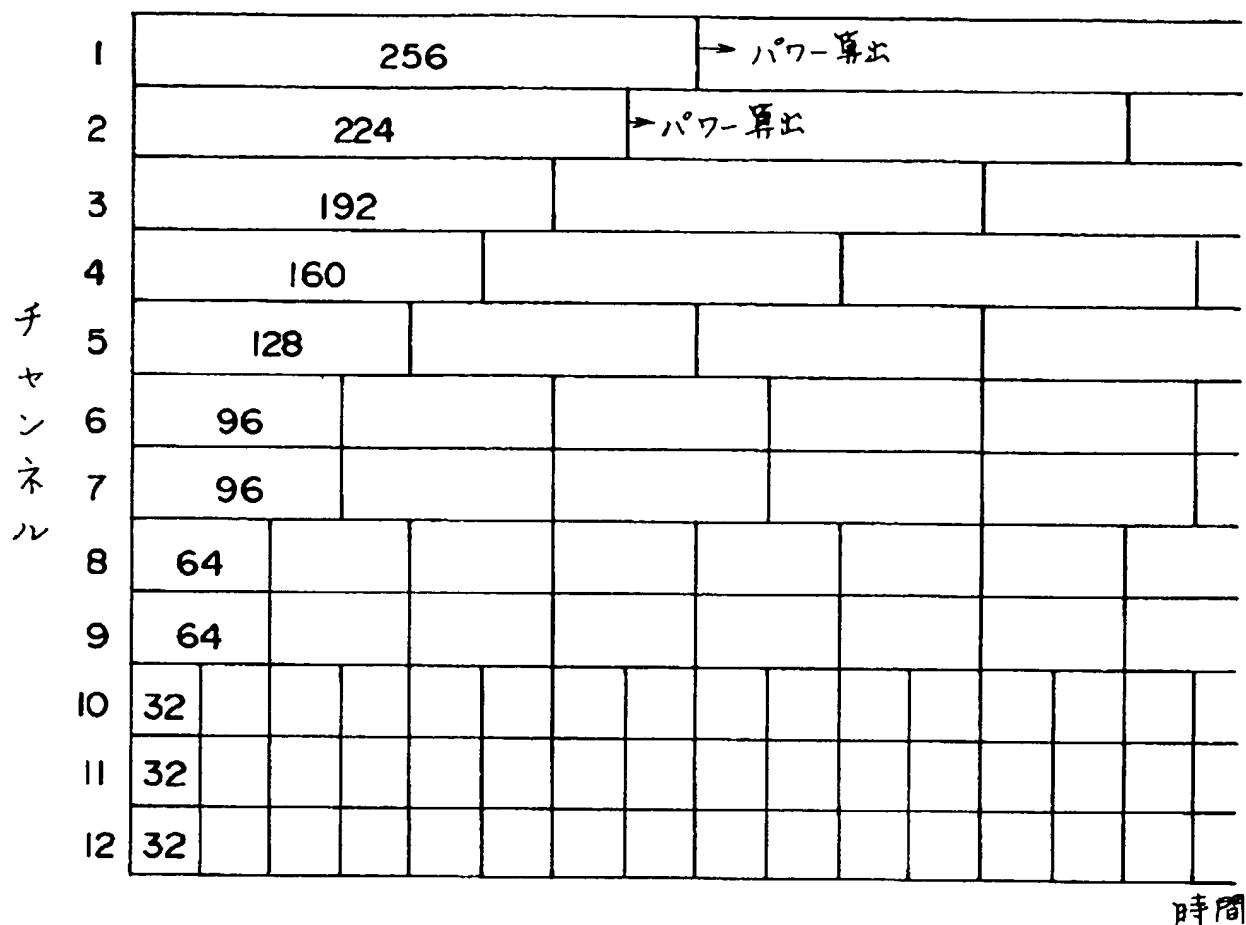
【図2】

CH 番号	通過帯域 (Hz)	帯域幅 (Hz)	帯域幅 の逆数比	時間 窓長 (サンプル)	パワー出力 周期 (サンプル)	データレート サンプル/秒
1	340-412	72	8.2500	512	256	31.25
2	412-499	87	6.8276	448	224	35.71
3	499-605	106	5.6038	384	192	41.67
4	605-733	128	4.6406	320	160	50
5	733-887	154	3.8571	256	128	62.5
6	887-1075	188	3.1596	192	96	83.33
7	1075-1303	228	2.6053	192	96	83.33
8	1303-1578	275	2.1600	128	64	125
9	1578-1912	334	1.7784	128	64	125
10	1912-2316	404	1.4703	64	32	250
11	2316-2806	490	1.2122	64	32	250
12	2806-3400	594	1.0000	64	32	250
計	—	—	—	—	—	1387.79

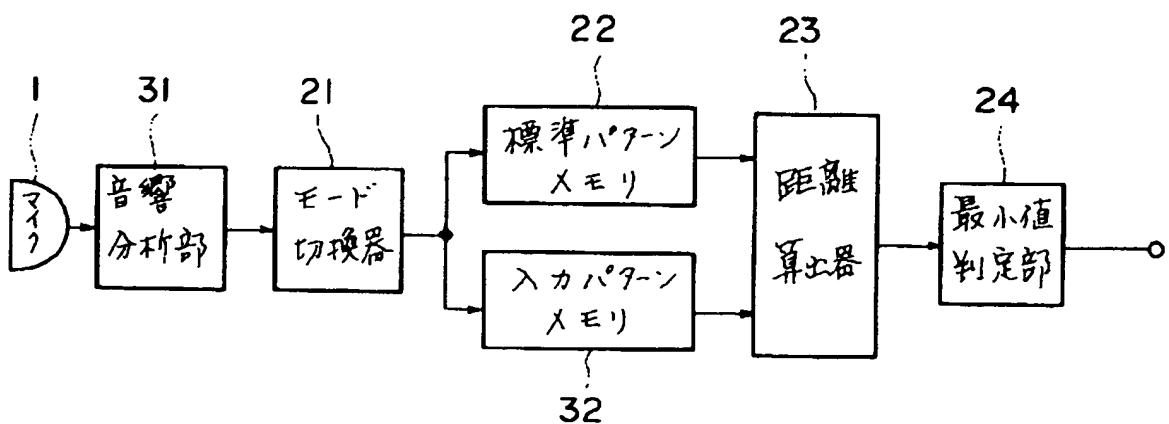
【図3】



【図4】



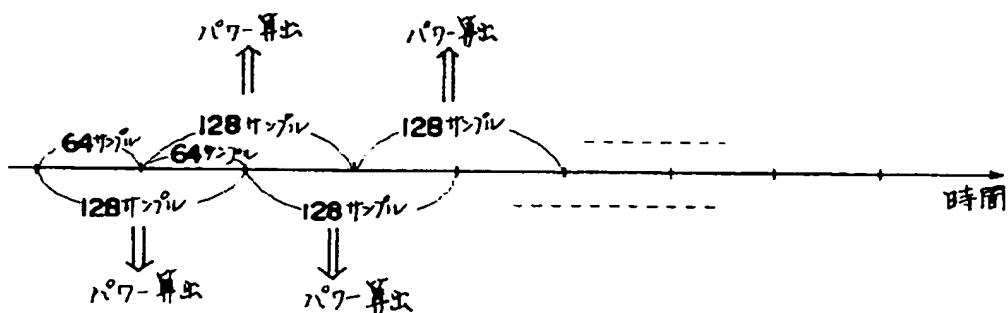
【図5】



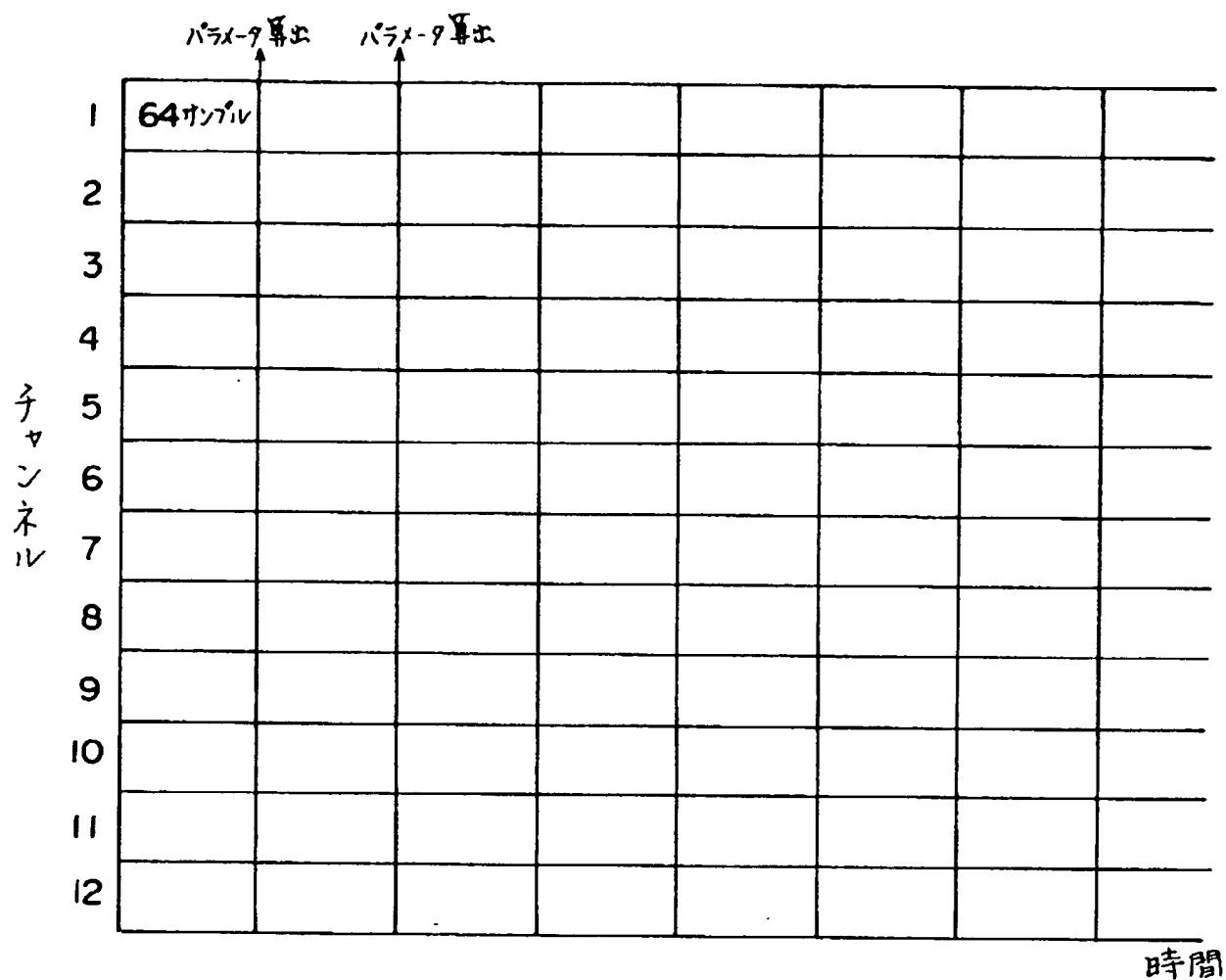
【図6】

CH番号	通過帯域[Hz]	バンド幅[Hz]
1	340-412	72
2	412-499	87
3	499-605	106
4	605-733	128
5	733-887	154
6	887-1075	188
7	1075-1303	228
8	1303-1578	275
9	1578-1912	334
10	1912-2316	404
11	2316-2806	490
12	2806-3400	594

【図7】



【図8】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.